DOCUMENT RESUME

ED 136 558 FL 008 354

AUTHOR Ross, Robert N.
TITLE Ellipsis and the Structure of Expectation. San Jose State Occasional Papers in Linguistics, Vol. 1, November, 1975.
INSTITUTION San Jose State Univ., Calif. Linguistics Program.
PUB DATE 75
NOTE 10p.; Paper presented at the Annual California Linguistics Conference (Fifth, May 3-4, 1975)
AVAILABLE FROM Robert N. Ross, Division of Psychiatry, Boston University School of Medicine, 720 Harrison Ave., Boston, Massachusetts 02118 ($1.00)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Connected Discourse; *Deep Structure; *Discourse Analysis; Grammar; Language Research; *Semantics; Sentences; *Sentence Structure; *Structural Analysis; Surface Structure; Syntax; Verbs
IDENTIFIERS *Ellipsis

ABSTRACT

This paper discusses one way of exploring how we perceive and understand the connections between some parts of texts, or between one sentence and the whole discourse. Understanding ellipsis involves non-syntactic understanding; the semantic structure is responsible for our understanding of elliptical sentences and encoding the knowledge contained in them. These covert pieces of information are referred to as "structures of expectation." The structures of expectation responsible for our comprehension of connected discourse can be demonstrated by showing how explicit arguments satisfying our expectations are supplied by information given later in the text. The approach was tested in an experiment in which several groups of people were given three sets of sentences with the structure: "A did X. B did Y." Subjects were asked to explain the situation in each set by drawing inferences about the relation between the two parts of each set. One clue to ways of relating the pairs of sentences is stress; another is argument-sharing. This type of analysis can teach us something about the structures underlying our inferences from the manifest content of texts. (Author/AM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED136558 —

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Ellipsis and the structure of expectation

Robert N. Ross
Boston University School of Medicine

I am interested in the semantics of texts larger than sentences; and I am especially interested in the semantics of stories. My interest is motivated by some quite practical concerns: for instance, we know the people around us chiefly on the basis of shared stories; anthropologists learn about the value systems of alien cultures by studying the stories they discover there; psychiatrists come to understand the emotional difficulties of their patients by making complex inferences from their patients' stories about themselves. The activity is commonplace; but it is nonetheless difficult. So it is well worth our while to look carefully into the ways in which information is encoded in stories.

But studying the semantics of stories is difficult because stories are more important for what they do not say than for what they in fact do say. Perhaps I can make this clearer by distinguishing stories from lists: in attending to a list, we deal literally and verbatim with its items; in attending to a story, we come to understand the connections between the items. And such connections between, of course, do not exist in any of the items taken singly.

So, to state my case briefly, I am interested in how we perceive and understand the connections between some parts of texts. Today I want to describe one possible strategy for studying this mysterious region: this is very much a statement of research in progress.

Discourses are not, as was once thought, mere assemblages of sentences such that in understanding the linguistics of the sentence we can also understand the linguistics of larger segments of speech. Rather we see now that the structure of the sentences themselves are closely related to the structure of the whole discourse in which they participate. Rules governing the use of proforms and other referentials are examples of what I mean. And further, recent developments in generative semantics are making even more subtle relations between sentence and discourse increasingly apparent.

I would go at this same material by studying how people deal with ellipsis. Ellipsis depends upon non-syntactic understanding of language. Take, for example, this little conversation: "Joe, the car." "Oh, no." "Oh yes." "Oh well."

There are no syntactic markers and yet we do understand a good bit of the situation represented in these words. In a sense, we understand the "shape of it," the configuration of its elements, even though we cannot be sure about the specific contents. That is, Joe could have forgotten the car; he could have parked it on a hill without securing the hand brake; and so forth. That kind of information we do not know. We do know, however, that something untoward is happening in which the car is involved, that someone else is calling it to Joe's attention, and so forth.

This is done without the aid of syntax. Instead of overt syntactic markers, there must be some very complex underlying network of semantic structures responsible for our comprehension of the scene. That is, this stretch of speech must be so constituted that we can readily imagine using it to describe a scene in the real world. And if that is so, someplace in this stretch of elliptical sentences we must be able to find where that common knowledge about the ways of the real world gets encoded.

FL008354

If some such covert pieces of information were not there, we would not have been able to fill in the "missing information" to make this set of words coherent and understandable. Those covert pieces of information are what I am calling structures of expectation. And I suspect they reflect some of the semantic structures we hold in common with our interlocutors when we communicate our stories.

So, again, it is not a trivial concern of ours to examine the process as closely as we can.

The case grammars developed by Fillmore and Chafe furnish useful analytic and descriptive tools for explaining how we comprehend discourse. We can describe the semantic structure of a text, for instance, by noting the relations between the explicitly stated cases and the implicit ones.

How we detect implied cases requires some comment: We can know what an implicitly stated argument is because we can know with some certainty what are the so-called diagnostic arguments -- the ones required by verbs as part of their definition -- and then too we can know the other arguments not required by the verb but only adding circumstantial detail. We can thus examine the relations existing between arguments bound to their verbs in the surface of the text and those arguments required by the verbs but not necessarily appearing in the text. In this way, we can demonstrate the structures of expectation responsible for our comprehension of connected discourse: we can show how explicit arguments satisfying our expectations are supplied by information given later in the text to fill otherwise "empty slots" marked as necessary by the verbs but not directly filled by their arguments.

This approach lends itself to direct experiment. For we can describe the semantic structure of sentences and then give them to people to see what they do with them.

I have given several different groups of people three sets of sentences with the instructions of "explain the situation." The sentences are:

(1) John broke the window. Jim found the rock in his living room.
(2) Arthur threw the ball into the woods. Barbara was very angry.
(3) Mary robbed the bank. Her mother had to go to the hospital.

Superficially, these three sets have similar structure: "A did X. B did Y." Because they are syntactically similar, it is important for us to note that people used quite different strategies for dealing with the various sets of sentences. Those strategies, then, must have been motivated by more semantic concerns.

The instructions were to "explain the situation"; and people were free to give any explanation they felt to be appropriate. Moreover, they could spin as long a story as they cared to about the situations they were explaining. But the interesting thing is that despite this apparent liberty, people's "explanations" fell into only a few characteristic groups.

As the data shows, for each of the three sets of sentences, there were only a few (or only one) popular answers accounting for the bulk of responses; and then there was a set of more or less idiosyncratic answers. Right now I am interested in the popular answers. But it should be mentioned that although there are "normal" answers, the other idiosyncratic ones cannot be called "wrong."

TABLE OF DATA: ELLIPSIS TEST

Response	% Giving* Response	M	F	% Adding** New Material
Test I. JOHN BROKE THE WINDOW. JIM FOUND THE ROCK IN HIS LIVING ROOM.				
A. John broke Jim's window.	77 (34)	19	15	35 (12)
B. No direct explanation.	14 (6)	1	5	83 (5)
C. Two separate events.	7 (3)	3	0	(1)
D. Jim broke his own window.	2 (1)	0	1	---
Test II. ARTHUR THREW THE BALL INTO THE WOODS. BARBARA WAS VERY ANGRY.				
A. It was Barbara's ball.	30 (13)	8	5	62 (8)
B. Arthur disrupted the game.	18 (8)	3	5	63 (5)
G. No direct explanation.	18 (8)	5	3	38 (3)
F. She did not want to chase it.	11 (5)	1	4	100 (5)
H. Two separate events.	7 (3)	1	2	---
E. She knew she was being lured into woods.	4 (2)	0	2	100 (2)
C. Ball hit her on the head.	4 (2)	1	1	100 (2)
D. Maternal discipline.	4 (2)	1	1	100 (2)
I. She could not find the ball.	2 (1)	0	1	---
Test III. MARY ROBBED THE BANK. HER MOTHER HAD TO GO TO THE HOSPITAL.				
A. Mary had to pay mother's hospital bills.	39 (17)	9	8	41 (7)
B. Mother had heart attack, breakdown, etc.	23 (10)	5	5	40 (4)
F. No direct explanation.	11 (5)	3	2	20 (1)
C. Mother goes to care for Mary.	9 (4)	2	2	100 (4)
D. Mother was hurt trying to stop Mary.	9 (4)	2	2	25 (1)
E. Mother was shot by police.	7 (3)	1	2	100 (3)
G. Two separate events.	2 (1)	1	0	100 (1)

*Number in parentheses is actual number of people giving response (n=44).
**This is per cent of those people giving the response who added new facts to rationalize their answers.

Just about everybody explained the sentences JOHN BROKE THE WINDOW. JIM FOUND THE ROCK IN HIS LIVING ROOM by saying that JOHN BROKE JIM'S WINDOW. About 3/4 of the people said exactly this; most of the others said this, but less definitely.

The second set of sentences ARTHUR THREW THE BALL INTO THE WOODS. BARBARA WAS VERY ANGRY had four fairly popular answers (accounting for 77% of the responses). The most popular (given by 30%) was IT WAS BARBARA'S BALL. And, as the table of data shows, smaller groups of people gave the other characteristic answers.

The third set of sentences MARY ROBBED THE BANK. HER MOTHER HAD TO GO TO THE HOSPITAL had three answers (accounting for 73% of the responses). By far the most popular was MARY ROBBED THE BANK TO GET THE MONEY SHE NEEDED TO PAY HER MOTHER'S HOSPITAL BILLS. A smaller group said that the mother HAD A HEART ATTACK OR A NERVOUS BREAKDOWN WHEN SHE HEARD THAT HER DAUGHTER HAD ROBBED THE BANK AND SO THE MOTHER HAD TO GO TO THE HOSPITAL.

This is the data. Now how do we explain (1) the grouping of these responses so that some few of them are fairly popular? and (2) how do we explain the fact that even though we do find groups of popular answers, the more idiosyncratic ones cannot be called "wrong"? For instance, one woman explained the first set of sentences by saying that JOHN BROKE HIS OWN

WINDOW AND JIM FOUND THE ROCK. She went on to say that she just hated to think about people destroying other people's property.

We can infer the sorts of semantic decisions people must have been making by studying how they got from information manifest in the text to information not explicitly given. And we should be able to make such inferences from the data I have just exhibited.

The first clue we might follow is that of contrastive stress in the two sentences of each set. For it is becoming clear that the relations between the stressed and the unstressed portions of sentences mark the relations between the new and the given information of those sentences. Or, if not exactly that, at least they mark the difference between items of content that should be attended to and others that can be more or less taken for granted.

On the basis of this simple observation, we can assert that the problem people faced in "explaining" these elliptical sentences was the problem of relating new (that is, stressed) items of information in the two sentences. Differences in interpretations of the two sentences, likewise, could be attributed to people's taking different items as important (that is, stressed). For instance, let us consider the various ways in which the sentences could be accented.

I. 1. JOHN broke the window.	1. JIM found the rock in his living room.
2. John BROKE the window.	2. Jim FOUND the rock in his living room.
3. John broke the WINDOW.	3. Jim found the ROCK in his living room.
	4. Jim found the rock in HIS living room.
	5. Jim found the rock in his LIVING ROOM.
II. 1. ARTHUR threw the ball into the woods.	1. BARBARA was very angry.
2. Arthur THREW the ball into the woods.	2. Barbara WAS very angry.
3. Arthur threw the BALL into the woods.	3. Barbara was VERY angry.
4. Arthur threw the ball into the WOODS.	4. Barbara was very ANGRY.
III. MARY robbed the bank.	1. Her MOTHER had to go to the hospital.
2. Mary ROBBED the bank.	2. Her mother HAD to go to the hospital.
3. Mary robbed the BANK.	3. Her mother had to GO to the hospital.
	4. Her mother had to go to the HOSPITAL.

Further research, using a tape recorder to capture the actual intonations, is of course necessary. And in that research, I would follow this clue: Of all the possible connections in the first set of sentences, it would appear that most people (77%) chose to relate WINDOW and ROCK. This is surely not the only possible response. For we could just as well relate BROKE and FOUND in a meaningful way. Or, more interesting, we could relate JOHN and HIS as the one woman in fact did. Meaning: John broke his own window and Jim found the rock.

There are many waysof relating the sentences in the second set, too. But most of the different "explanations" (65%) were attempts to relate BALL and BARBARA. It would appear that ARTHUR and ANGRY also received stress. This accounts for the bulk of answers: A, B, C, F, and I. A few people, however, focused their attention on the WOODS. They answered by saying that BARBARA KNEW SHE WAS BEING LURED INTO THE WOODS AND SO SHE WAS ANGRY. Another small group tried to relate ARTHUR and BARBARA by saying that SHE IS HIS MOTHER AND SHE IS ANGRY BECAUSE HE THREW HIS BALL INTO THE WOODS.

Of the many possible ways of relating the pairs of sentences in the third set, likewise, only a few are actually chosen. The first group of responses (39%) represents people's attempt to relate BANK and HOSPITAL. That is answer A. Another group of responses (39%) represents people's attempts to relate ROB and MOTHER. That is, answers B,D, and E. And, finally, another small group attempted to relate MARY and MOTHER (9%) by saying that THE MOTHER IS GOING TO THE HOSPITAL TO CARE FOR MARY. That is, answer C.

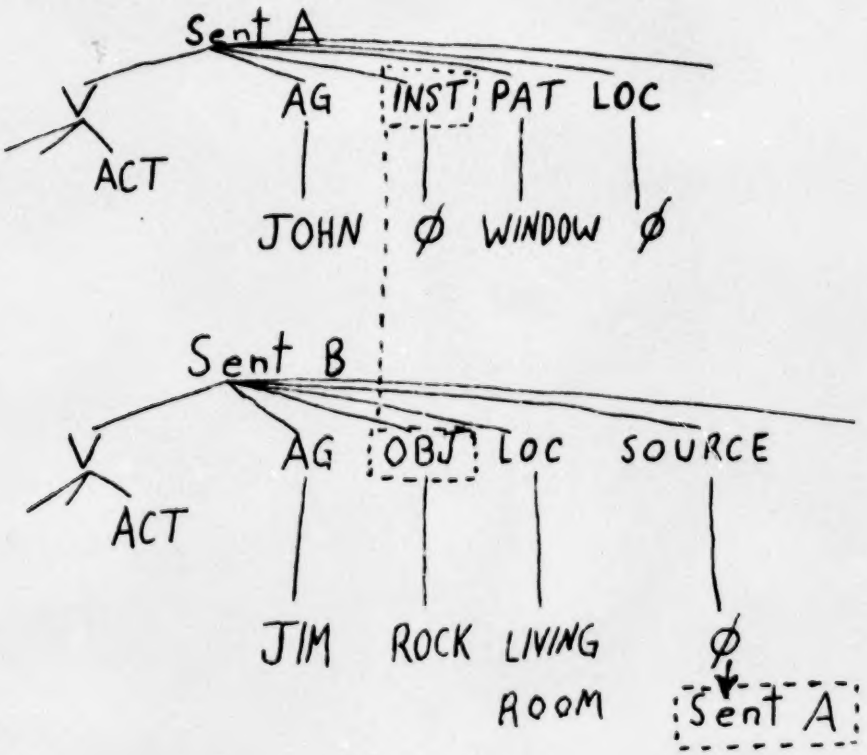
These differential stresses are the signals we readily understand and work with in trying to make sense of these pairs of elliptical sentences. They are the phonological markers indicating something of the underlying semantic structure holding the sentences together in people's comprehension and, presumably, in their production as well. They are signals that work independent of syntax, too.

But although they mark what is important in the sentences and they mark what items of content can be connected with what, they do not indicate how the items are important. For that kind of information, we must turn to a more content-based description of how the sentences work and how people can attend to their meanings.

I propose explaining the semantic behavior I have just described by using the notion of covert cases and other semantic descriptors made available to us through the work of Chafe, Fillmore, and more recently several others.

Now let's get down to cases.

We can show the case relations in the first pair of sentences as follows:



There are probably many subtle factors responsible for our being able to "put the two sentences together"; but one easily described factor is argument-sharing. By this I mean that a nominal argument bound to one verb serves as a nominal argument bound to another verb as well. This sort of argument-sharing reflects one way of binding the two sentences together in structures of inference.

The verb BREAK, for instance, requires three nominal arguments: an AGENT to do the breaking; an INSTRUMENT by which to do the breaking; and a PATIENT to get broken. In the test sentence, the INSTRUMENT required by the verb BREAK is not given.

The second sentence, however, has information that can fill the empty slot in the first sentence. The verb FIND requires several arguments: an AGENT to do the finding; an OBJECT to get found; and a SOURCE of the found object.

These arguments are diagnostic: they tell what the verb means.

If we look at the two sentences with this set of descriptive labels in mind, we learn several things about how we come to understand them. For one thing, we make the tactical assumption that the LOCATION of the first sentence is pretty much the LOCATION of the second; and that the WINDOW broken is part of the-LIVING ROOM in which the ROCK is found.

More to the point, however, is the way in which we link the INSTRUMENT required by the verb BREAK and the OBJECT required by the verb FIND; and we rather automatically assume that the rock found is the instrument implied and not stated in the first sentence.

We also understand that the verb FIND requires a SOURCE of the thing found. And so we make the whole of the first sentence into the SOURCE of the OBJECT found in the second sentence.

Being a SOURCE, the first sentence is said to entail the second: that is, because of sentence A, then sentence B.

This is a straightforward case; and 3/4 of the people responded in just this way to relate the broken window and the rock.

The second set of sentences is a more complicated instance of the same argument-sharing. Remember that some 65% of the people found that the problem was that of relating BALL and BARBARA and to account for ARTHUR and ANGRY in some secondary way. We can diagram these two sentences as we did the first set and make similar kinds of arguments to explain how people came to give the responses they did.

The verb THROW requires the arguments: an AGENT to do the throwing; an OBJECT to get thrown; and a SOURCE of the OBJECT. That much is easy. The second sentence, however, contains no information about the ball; and yet 30% of the people said IT WAS HER BALL. Where did they get this information?

If we look more closely at that second sentence, we see that the verb BE ANGRY requires several arguments: an EXPERIENCER to feel the anger; an INSTRUMENT or some event that produces the anger; and a SOURCE of the anger or some situation or evaluation that endows the INSTRUMENT with its efficacy. This is not entirely clear. But simply throwing a ball does not necessarily make someone angry. In understanding this pair of sentences, however, we assume that Barbara gets angry because Arthur threw the ball and because his throwing the ball is a meaningful act.

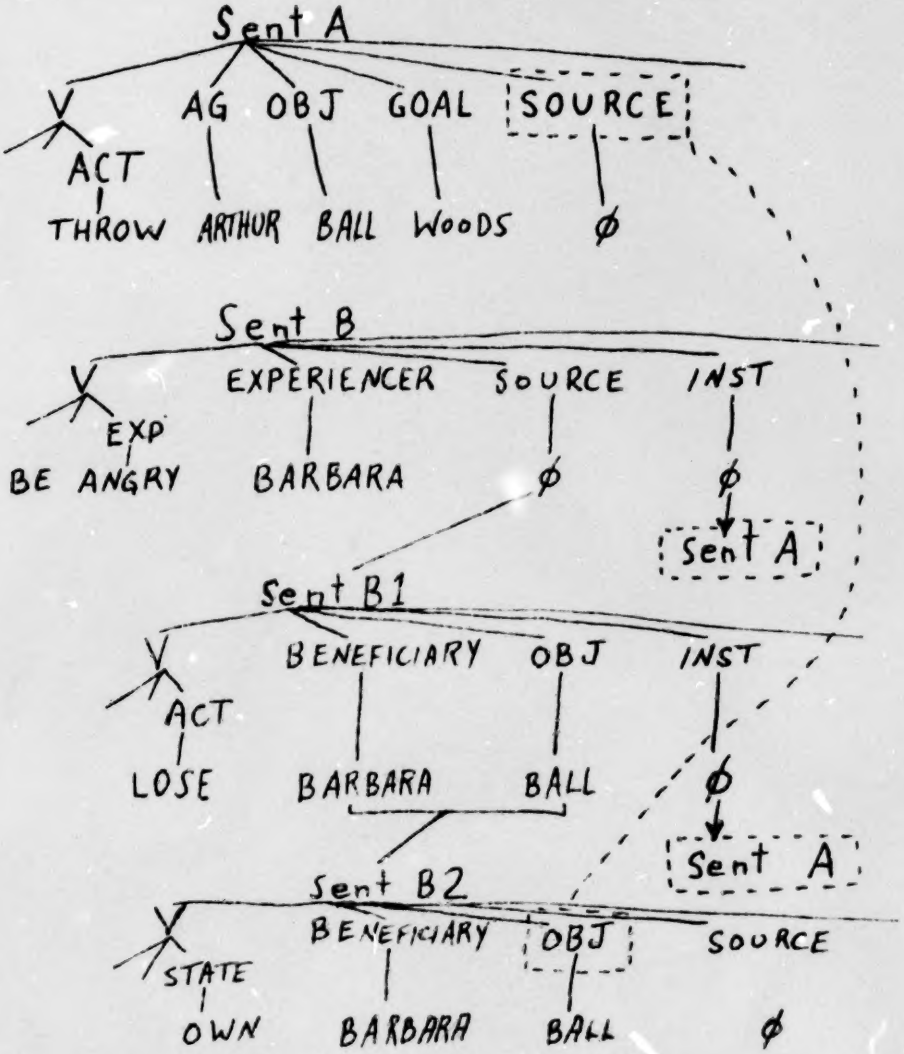
We find that the INSTRUMENT of Barbara's anger is stated as the whole of the first sentence: ARTHUR THREW THE BALL INTO THE WOODS. The SOURCE of her anger is more of a semantic problem. And different people solved the problem differently.

The SOURCE of Barbara's anger for the 30% of the people who said IT WAS HER BALL is that she lost her ball or was, at the veryleast temporarily deprived of it. So that embedded in the sentence BARBARA WAS VERY ANGRY must be another sentence related to the first sentence and related, at the same time, to the sentence BARBARA WAS VERY ANGRY.

That is to say, the context "creates" a sentence BARBARA LOST A BALL. And the INSTRUMENT of that loss, as we have already seen, is the INSTRUMENT of her anger: ARTHUR THREW THE BALL INTO THE WOODS.

Embedded in this already embedded sentence is yet another: BARBARA OWNS A BALL. And, of course, among the arguments required by the verb TO OWN is an OBJECT.

Now back to the first sentence where we have already said that THROW requires the argument SOURCE. What you throw must come from someplace. The various arguments of these several sentences just outlined fit together in such a way that the SOURCE of the ball thrown in the first sentence is the OBJECT implied in the second sentence: BARBARA OWNS A BALL. And so, 30% of the people gave as their "explanation" that BARBARA WAS ANGRY WHEN ARTHUR THREW THE BALL INTO THE WOODS BECAUSE IT WAS HER BALL. Actually, they did not say all this at all. They simply said IT WAS HER BALL and easily expected us to understand the rest of the context.



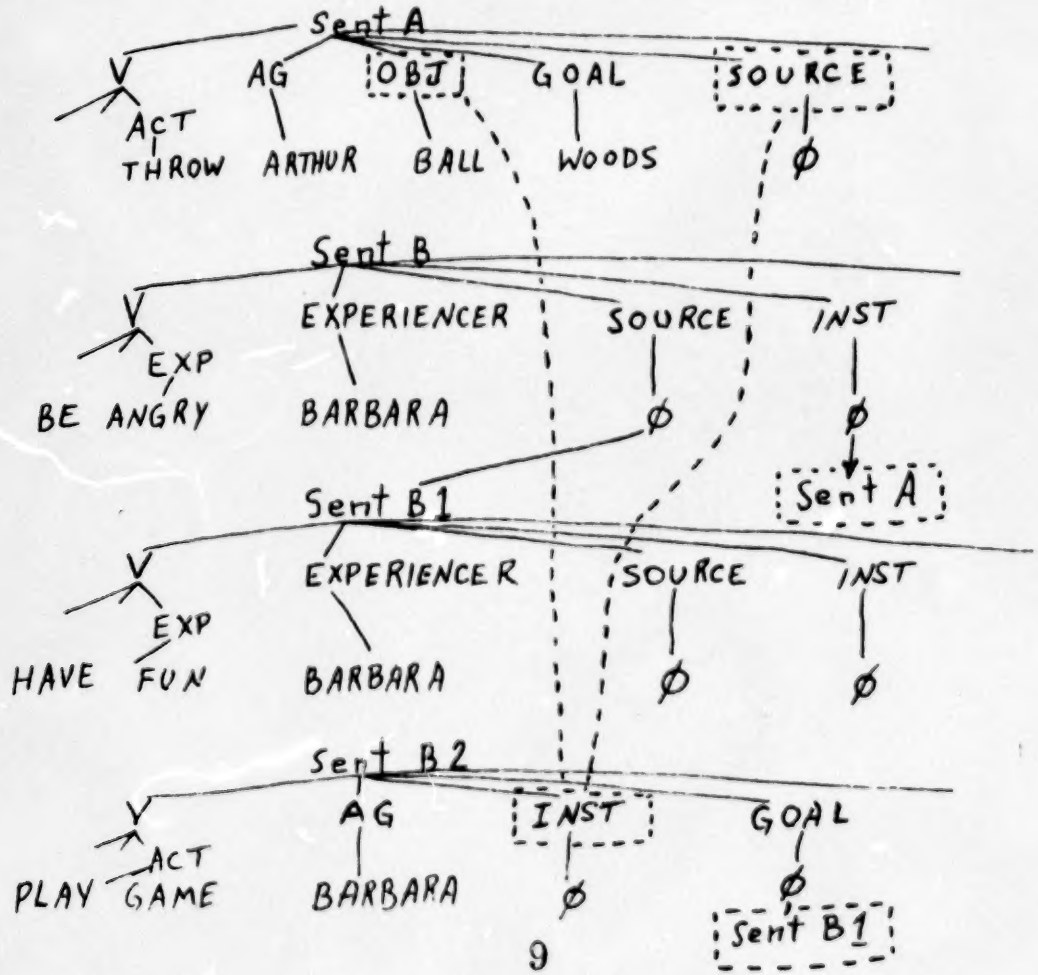
We can contrast this response with the somewhat less popular one that ARTHUR DISRUPTED THE GAME SHE WAS PLAYING WITH THE BALL. For in this instance, the ways in which the arguments get shared and the embedded sentences "created" by the context are quite different.

The first sentence of the pair will, of course, we diagrammed as it was in the previous example. The difference between this response and the one I have just discussed lies, rather, in the different ways of treating the second sentence: BARBARA WAS VERY ANGRY.

As in the previous example, the INSTRUMENT of her anger, that is what makes Barbara angry, is that ARTHUR THREW THE BALL INTO THE WOODS. The whole of the first sentence is the INSTRUMENT. The SOURCE of her anger, however, is different.

The SOURCE of Barbara's being angry is given in some embedded sentence like: BARBARA WAS HAVING FUN. The verb TO HAVE FUN requires among its various arguments, naturally, some INSTRUMENT by which the fun is accomplished. This INSTRUMENT is given, in this instance, by another embedded sentence: BARBARA WAS PLAYING A GAME WITH HER BALL.

Now the verb THROW in the first sentence, as we have seen, requires a SOURCE; and the SOURCE is given in this covert argument implied in the second sentence as the INSTRUMENT OF HER GAME. And so we can explain the responses of those 18% of the people who said that ARTHUR DISRUPTED THE GAME AND MADE BARBARA ANGRY.



Since I am primarily interested in showing the kinds of explanations we can give, and not to exhaust the possibilities, there is little point in going through the third set of sentences.

In closing then I would like to mention what I consider to be the value of this sort of work.

First of all, by inquiring into the arguments we bind to the various verbs, we are in effect creating a list of semantically important distinctions in our culture. An AGENT is not a PATIENT; a BENEFICIARY is not a RECIPIENT; and so forth. This sort of research is useful in itself. But it also allows us to understand what we mean by "reading between the lines." For what exists "between the lines" must look something like what I have been outlining here.

And if we can, in fact, discover such structures of language underlying the inferences we can draw from the manifest contents of texts, then we can also learn a great deal about the usual and the unusual ways people have, in general, of drawing inferences from their verbally encoded experience. And maybe we can even learn something about the more pathological ways of encoding and understanding experience.

In short, perhaps we can learn something about how our language, our meanings, and our behaviors are related.